

本周工作

(解聪 2013.9.23-9.30)

1. 主要完成了 MUIE 论文的修改。视频还在制作中
2. 最近三天和尚平平讨论了项目。讨论内容包括:
 1. 项目的任务:目前看来还是挖掘用户群之间,用户行为之间以及用户群与用户行为之间的关联。(至于之前说的规则的挖掘,我的理解是规则是用来分类的,这样的话任务就变为了对用户群或行为进行分类。)
 2. 可视分析可以做的:因为自动关联性分析方法很完善很全面。我们不大可能提出新的算法,也不太可能大幅度提高现有算法的计算效率(这虽然是大数据中需要解决的问题,但是我们的数据集不大,而且很难找到可是分析的切入点)。我们能做的就是改进现有算法在特定应用场景下的不足。
 3. 如何做:最大可能还是使用我们现在已经有的几个数据,从中发现案例。现在比较好的是淘宝的类目关联的案例,但是我们的结果还是比较粗粒度,我们应该先尽量在这个案例上进一步深入分析。
 4. 具体实施方案:利用交互修正机器自动方法不准确的地方。(比如:机器算类目关联的时候考虑了所有人群,分析师可以在像素图上选择选择,忽略掉一些人群来做关联分析。)交互应该主要针对像素图来做易发挥人眼对图像的优势;投影部分 MDS 的慢而且结果不好,现在再换为 SOM 试试。

下周工作

下周主要精力投入在用户群行为关联分析项目中。现有的淘宝的案例只是基于统计信息的可视化,我们需要进行细粒度并且包含更高维度的分析。